THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 62, No. 3, March 1983
Printed in U.S.A.

Variable Rate ADPCM Based on Explicit Noise Coding

By N. S. JAYANT

(Manuscript received July 15, 1982)

This paper discusses a variable bit rate speech coding system based on explicit coding of the reconstruction noise in ADPCM (differential pulse code modulation with adaptive quantization). If the ADPCM bit rate is R *bits/sample, PCM coding of its noise using an average bit rate of* R_n *bits/sample provides the receiver with the possibility of operating at any bit rate in the range* R *to* $R + \max\{R_n\}$. *Using* R *values in the range 2 to 5, and* R_n *values in the range 0 to 3, we compare the performance of the* $(R + R_n)$*-bit system with that of conventional* $(R + R_n)$*-bit ADPCM. If noise coding is based on instantaneous* R_n*-bit quantization of its samples with an optimized step size, the signal-to-noise ratio performance is comparable to that of conventional ADPCM for* $R_n = 1$*, but it deteriorates significantly for* $R_n > 1$*. With non-instantaneous noise coding, the performance can exceed that of conventional ADPCM for any* $R_n > 1$*, if* $R > 2$*. This is due to a variable bit allocation algorithm that quantizes noise samples with differing resolutions, while maintaining a constant total bit rate in every block of 4 ms. The algorithm does not require the transmission of any extra side information. It can also be regarded as a way of improving the performance of ADPCM coding at a single bit rate of* $R + R_n$ *bits/sample.*

I. INTRODUCTION

Multiple-stage coding, where the reconstruction noise from an initial stage is itself coded for transmission in a subsequent stage, is known to provide substantial gains over single-stage coding in the context of deltamodulation using oversampled inputs.[1,2] In this paper, we consider two-stage systems for multibit differential pulse code modulation with adaptive quantization (ADPCM) coding of Nyquist-sampled speech inputs. Unlike systems that permit oversampling, signal-to-noise ratio

(s/n) gains in our systems will be seen to be either slightly negative, or positive but nondramatic. However, the proposed systems have a feature that is common to all noise-coding systems, the property of embedded coding: the output bit sequence of the coder contains a subsequence that can be used in a straightforward manner to provide lower bit-rate operation with an output speech quality very close to that of conventional operation at the lower bit rate; as a result, the channel or receiver can switch, as needed, between low-rate and high-rate modes. The possibility of *variable-rate* operation is a very desirable feature in digital communication systems such as packet-switched voice networks.[3] A PCM coder is inherently an embedded coder. Least significant bits in a PCM codeword can be progressively dropped, with a graceful loss of quality that is no greater than about 6 dB/bit. Conventional differential pulse code modulation (DPCM) is not an embedded coding system in a similar sense because of the presence of a feedback loop in coder and decoder.

Explicit noise coding is not the only way of designing an embedded ADPCM system. Coarse feedback in the DPCM predictor loop[4,5] is known to provide a very robust basis for embedded DPCM, with very little s/n degradation compared to conventional DPCM at a given bit rate; and the results are also expected to extend to DPCM with an adaptive quantizer. In the coarse-feedback approach, the encoder performs an appropriate quantization of predictor input in anticipation of a similar quantization that may be forced at the receiver as a result of bit-dropping. The coarse feedback embedded system can also drop more than one bit, in a progressive fashion, to provide a wide range of bit rates. The noise-coding approach provides zero degradation of quality at the lower bit rate, R. More important, explicit noise coding offers the possibility of complex versions of $(R + 1)$-bit ADPCM that can provide positive performance gains over conventional $(R + 1)$-bit coding. ADPCM with variable bit allocation (Section V) is one example of such a complex system. The noise-coding system with variable bit allocation can also be used as a *single-rate* coder in which the coding process is split into two steps (conventional ADPCM followed by noise coding) to permit a simple form of time-domain bit allocation for the improvement of ADPCM performance.

Figure 1 is a block diagram of a variable-rate coder employing optional coding of ADPCM noise, with an average noise-coding rate of R_n bits/sample. The special case of $R_n = 1$ is treated at length in Sections IV through VII. When the dashed boxes for bit allocation are eliminated, instantaneous noise-coding results, with a coding rate of exactly R_n bits for every noise sample. When the parts of the system within boxes A or B are eliminated, $R_n = 0$, and conventional single-rate ADPCM results, with a total bit rate of R bits/sample. The

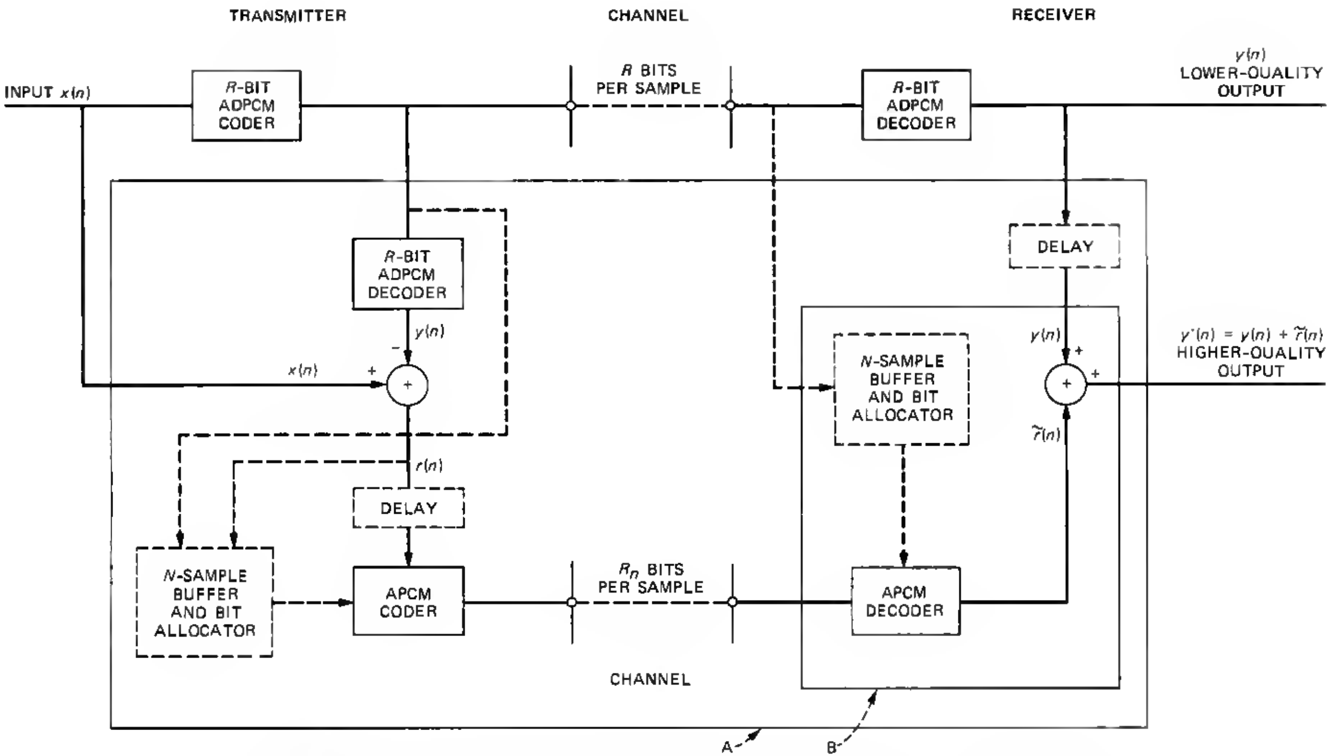


Fig. 1—Block diagram of a variable-rate ADPCM system. Variation of the average noise-coding bit rate R_n results in a variable-rate system with a total bit rate that ranges from R to $R + \max\{R_n\}$ bits/sample.

extreme upper part of the figure (outside of box A) shows a conventional R-bit ADPCM coder-decoder.[6] The rest of the diagram (the part included in box A) shows the blocks that perform R_n bit/sample coding of the reconstruction noise samples

$$r(n) = x(n) - y(n), \tag{1a}$$

where $x(n)$ and $y(n)$ are the input and the output of an R-bit/sample ADPCM system. When the system of Fig. 1 includes noise coding, the final decoded value is $y'(n)$, a refinement of the conventional value $y(n)$:

$$y'(n) = y(n) + \tilde{r}(n). \tag{1b}$$

The total bit rate of the system in Fig. 1 is $R + R_n$ bits/sample. Variable-rate coding results from the use of different values of R_n. Examples in this paper cover the range of $0 \leq R_n \leq 3$. The case of $R_n = 1$ is discussed at length before generalization to $R_n > 1$. With $R_n = 1$, the variable-rate system of Fig. 1 reduces to a dual-rate system, with a total bit rate of either R or $R + 1$ bits/sample.

The noise information can be altogether eliminated by the system ($R_n = 0$) to provide conventional R-bit operation. Alternatively, the noise information may be eliminated, as necessary, by the channel or receiver. If the receiver does the elimination, the part of the system that is eliminated is that within box B.

The results of this paper are based on simulations with three sentence-length utterances: "The *c*hairman cast three votes" (*f*emale speaker); "A *l*athe is a big tool" (*f*emale speaker); and "A *l*athe is a big tool" (*m*ale speaker). These speech inputs are identified in the rest of this paper as *CF*, *LF*, and *LM*. All inputs are band-limited to the frequency range 200 to 3200 Hz.

II. SUMMARY OF THE ADPCM AND APCM CODERS

The ADPCM coder in this paper uses first-order prediction with a time-invariant prediction coefficient of 0.85. It also uses an adaptive quantizer with a one-word memory.[7] As Fig. 2 shows for the examples of $R = 2$ and $R = 3$, the (uniform mid-rise) quantization characteristic $Q(x)$ is multiplicatively expanded or compressed at every sampling instant by a factor (step-size multiplier M) that depends only on the magnitude of the most recent quantizer output $y(n-1)$. If $\Delta(n)$ is the quantizer step size at time n,

$$\Delta(n) = M(|y(n-1)|) \cdot \Delta(n-1). \tag{2}$$

The function M takes on one of 2^{R-1} values in R-bit ADPCM. Recommended multiplier sets for $R = 2$ and $R = 3$ are included in Fig. 2. Recommended multipliers for $R = 4$ and $R = 5$ are tabulated in Ref.

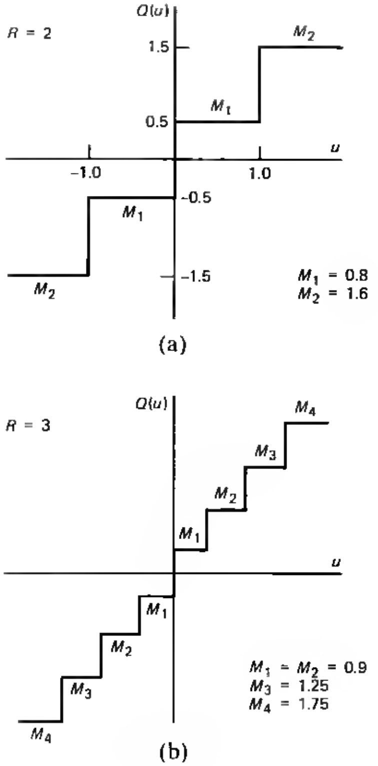


Fig. 2—Step-size multipliers used in (a) 2-bit and (b) 3-bit adaptive quantizers.

7. In the examples of Fig. 2, the use of the largest step-size multiplier also indicates the use of the outermost quantizer levels.

The adaptive PCM (APCM) coder for noise samples $r(n)$ will be described in detail in Sections IV, V, and VIII. The adaptive step size for the APCM coder will be seen to follow that of the quantizer in R-bit ADPCM. The purpose of the N-sample buffers in Fig. 1 is to permit a variable-bit-allocation procedure (Sections V and VIII) that provides a higher quality of noise quantization than what is possible with instantaneous quantization, the case of $N = 1$ (Section IV). When variable-bit allocation is employed, R_n will be interpreted as the *average* bit rate for noise coding. But the total number of noise-coding bits will be guaranteed to be a constant value, NR_n, for every block of N noise samples. The variable-bit allocation is first explained for the case of $R_n = 1$, implying noise coding with an average bit rate of 1 bit/sample (Section V). Extension to the case of $R_n > 1$ is straightforward (Section VIII).

III. RECONSTRUCTION ERROR $r(n)$

Figure 3a is a 16-ms-long speech segment from *CF*, and Fig. 3b illustrates the reconstruction-error waveform $r(n)$ in ADPCM for the example of $R = 4$. An important property is that $r(n)$ has occasional impulsive components. These are the slope-overload error bursts typical of DPCM with non-adaptive prediction. The extent of slope overload increases with coarseness of quantization. But, as seen in Fig. 3b, slope overload is quite evident even with $R = 4$ and adaptive quantization. In voiced speech, the time separation between slope overload bursts corresponds very closely to the pitch period. In the example of Fig. 3, this separation is about 40 samples (at 8 kHz), corresponding to a pitch period of about 200 Hz. Figures 3c and 3d will be discussed in Sections IV and V.

During slope overload, the noise samples $r_O(n)$ will have magnitudes in the range

$$0 \leq |r_O(n)| < \infty. \tag{3}$$

The limit ∞ can be replaced by a more meaningful finite value if the input is bounded, as in band-limited speech. But this won't be necessary for the purposes of this paper.

The non-impulsive background in the $r(n)$ waveform is associated with input samples that do not cause slope overload. In this granular noise region, the maximum magnitude of noise sample $r_g(n)$ is simply half the ADPCM step size:

$$0 \leq |r_g(n)| \leq \Delta(n)/2. \tag{4}$$

IV. THE $(R + 1)$-BIT CODER WITH INSTANTANEOUS ONE-BIT QUANTIZATION OF $r(n)$

From the theory of one-bit quantization, the reconstruction level $\delta(n)$ that provides the minimum mean square error with a one-bit noise quantizer is given by the mean absolute value of quantizer input:

$$\delta(n)_{\text{opt}} = E[|r(n)|]. \tag{5}$$

Ignoring slope-overload samples $r_O(n)$, and assuming that the magnitudes of the $r_g(n)$ samples are uniformly distributed in the range 0 to $\Delta(n)/2$,

$$\delta(n)_{\text{opt}} \sim E[|r_g(n)|] = \frac{1}{2} \cdot \frac{\Delta(n)}{2} = \frac{\Delta(n)}{4}. \tag{6}$$

Simulations have shown that the probability of slope overload is small enough for the above design to be indeed very close to the optimum. This is illustrated by the s/n versus $\delta(n)$ plots in Fig. 4 for $R = 4$, 3, and 2 bits/sample. The signal-to-noise ratio is maximum

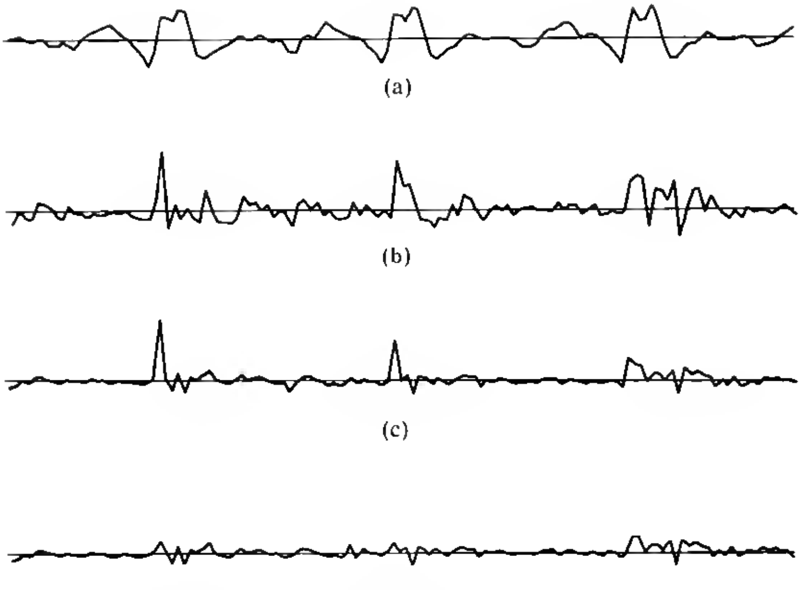


Fig. 3—(a) Input speech $x(n)$ (taken from CF) and reconstruction-error waveforms (b) (c) (d) in three ADPCM systems. All waveforms are 128 samples (16-ms) long. All error amplitudes are magnified by a factor of 10.

when the reconstruction level magnitude $\delta(n)$ of the 1-bit noise quantizer equals one-fourth the corresponding step-size $\Delta(n)$ in the R-bit ADPCM coder. When $\delta(n) = 0$, the system degenerates to the original R-bit ADPCM. Values at $\delta(n) = 0$ show the s/n of R-bit ADPCM.

Figure 3c shows the residual error after the $r(n)$ waveform (Fig. 3b) has been instantaneously quantized with a 1-bit/sample quantizer with reconstruction levels of $\pm\delta(n)_{\text{opt}}$. Note that the granular background components in $r(n)$ are uniformly reduced, but slope overload components are not.

The above step-size design implies that the noise coder is an instantaneous adaptive PCM (APCM) device that derives its step size from information that is already available in the R-bit ADPCM part of the $(R + 1)$-bit system. The N-sample buffer in Fig. 1 is not necessary for the operation of the instantaneous APCM coder.

The performance of the $(R + 1)$-bit system with instantaneous quantization is discussed at length in Section VI.

V. THE $(R + 1)$-BIT CODER WITH NON-INSTANTANEOUS ONE-BIT QUANTIZATION OF $r(n)$

Elimination of the impulsive components in $r(n)$ requires finer quantization. We now propose an algorithm that indeed allocates

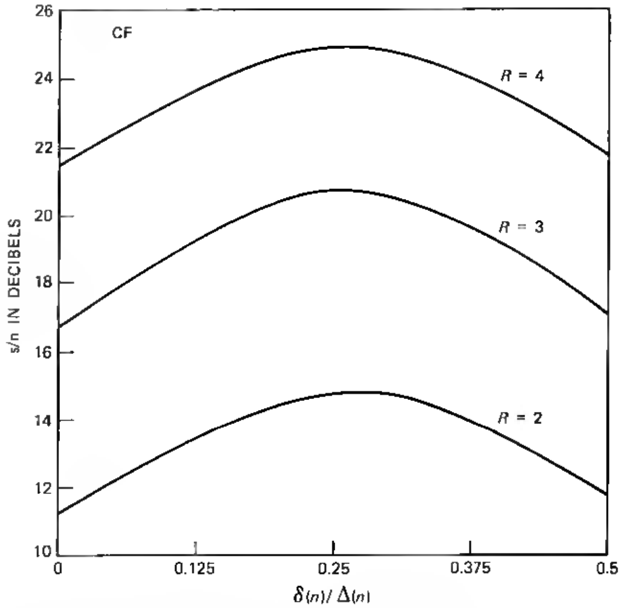


Fig. 4—Signal-to-noise ratio of $(R + 1)$-bit ADPCM system with instantaneous 1-bit quantization of reconstruction noise $r(n)$ from R-bit ADPCM.

$R_O > 1$ bit for slope-overload components, but still maintains an average bit rate of exactly 1 bit/sample in every block of N samples of $r(n)$. This is accomplished by assigning $R_g = 0$ bit/sample for granular noise components of very low magnitude in the block. The N-sample buffers and N-sample delays in Fig. 1 will be used to effect the above variable-bit assignment.

The location of slope-overload noise samples $r_O(n)$ and that of the low-magnitude granular noise samples $r_g(n)$ are both based on information that is already available to the R-bit ADPCM receiver, and therefore require no further side information to be transmitted.

The slope-overload samples are determined as those for which the quantizer output in R-bit ADPCM reaches the highest possible values for the given value of R (for example, levels associated with multiplier M_2 with $R = 2$ and levels associated with multiplier M_4 with $R = 4$; see Fig. 2).

The low-magnitude granular noise samples are located by rank-ordering $\Delta(n)$ values in tbe N-block, and by assigning zero bits to as many of these samples as necessary, in order of increasing $\Delta(n)$, until the total number of bits in the block is exactly N. While picking these zero-bit samples, it is very important to exclude samples associated with the use of highest output level. This precaution is needed because slope-overload errors can be associated with small values of $\Delta(n)$ as

well as large ones. In fact, as mentioned in the last paragraph, a defining cue for slope overload is not the value of the current *step size*, but rather the value of the current ADPCM quantizer *output level* (or current step-size *multiplier* if output levels and multiplier values have a one-to-one mapping, as in Fig. 2).

The net result of the above procedure will be to assign $R_O > 1$ bits where noise magnitudes are guaranteed to be the highest, and to assign $R_g = 0$ bits where noise magnitudes are guaranteed to be the smallest. The remaining samples are assigned 1 bit/sample as in Section IV.

The variable-bit-rate algorithm follows the constraint that the total number of bits per block is N:

$$N_O \cdot R_O + (N - N_O - N_g) \cdot 1 + N_g \cdot 0 = N; \qquad N_g = N_O(R_O - 1), \quad (7)$$

where N_O and N_g are the numbers of slope-overload and low-magnitude granular samples in N samples of $r(n)$. Note that the constraint above also implies that

$$N_O \cdot R_O = N_O + N_g \leq N; \qquad N_O \leq N/R_O. \quad (8)$$

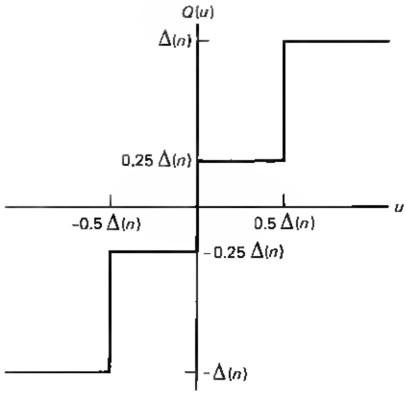
This latter constraint on N_O is explicitly enforced even in those cases where the number of maximum multiplier samples may exceed N/R_O, for a chosen R_O.

The design of R_O should reflect the probability of use of the maximum reconstruction level in the R-bit ADPCM coder. This probability controls the fraction N_O/N. As shown in Ref 7, this probability is a decreasing function of R; consequently, the maximum allowable value of R_O that does not violate (8) is an increasing function of R. In fact, in our experiments, we have found that for N values of interest, the s/n maximizing values of R_O happen to be very close to the number of bits/sample R in the basic ADPCM coder. Thus, for example, the slope-overload bursts in 3-bit ADPCM are quantized with a second stage of APCM coding with an appropriately designed 3-bit quantizer.

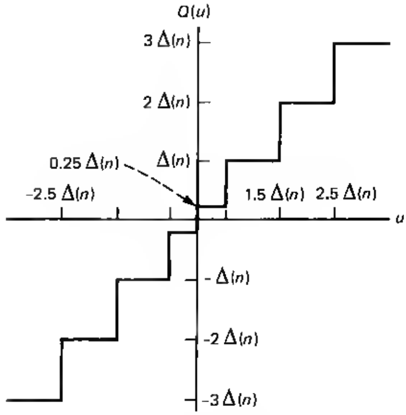
5.1 Design of noise-quantizer characteristic

Figure 5 illustrates quantizer characteristics that were experimentally found to provide nearly minimum mean square error in noise quantization. The smallest outputs in each of these characteristics are the $\pm\Delta(n)/4$ levels used in the instantaneous noise quantizer of Section IV. The largest output levels are $\pm\Delta(n)$ and $\pm 3\Delta(n)$ in the non-instantaneous quantizers for $R = 2$ and 3. For $R = 4$, the largest output levels in the noise quantizer will be $\pm 7\Delta(n)$. All these numbers obviously depend only on $\Delta(n)$, a value already available to the R-bit ADPCM receiver.

In one experiment with $N = 32$ and input *CF*, the number N_O of $r(n)$ samples coded with $R_O > 1$ bits/sample were 2, 3, 9, and 15,



(a)



(b)

Fig. 5—Quantization characteristics used for overload noise samples $r_O(n)$ in the non-instantaneous coding of ADPCM noise when the average rate for noise coding is R_n = 1 bit/sample. The ADPCM bit rate R is 2 in (a) and 3 in (b).

respectively, with 5, 4, 3, and 2-bit DPCM. These numbers reflect the much higher probability of using the maximum quantizer output level as R decreases. With the recommended design $R_O = R$, note that $N_O \cdot R_O < N = 32$ in all the four examples above, as required in (8). With $N = 128$ and the same input CF, values of N_O were 5, 9, 19, and 32, respectively.

5.2 Design of block length N

The buffer length N should be large enough so that for every noise sample coded with $R_O > 1$, there is an adequate selection of noise samples for which bit stealing ($R_g = 0$) is appropriate. However, the

quasi-periodic nature of slope-overload bursts (Fig. 3b) indicates that N need be no greater than the pitch period. This is indeed demonstrated in Fig. 6, which plots the signal-to-noise ratio of the $(R + 1)$-bit system as a function of N. Note that the performance at $N = 32$ (which is close to the pitch period 40 in Fig. 3 for CF) is very close to that at $N = 128$. Note also that the gain with $N = 32$, over the instantaneous quantization scheme of Section IV (the case of $N = 1$), is over 2 dB. Gains over $N = 1$ are less in the case of $R = 2$.

Figure 3d shows the residual error after $r(n)$ has been quantized with an average rate of 1 bit/sample, with $N = 32$ (a buffer length of 4 ms, with 8-kHz samples). Note that unlike the instantaneous quantization scheme of Fig. 3c, even the impulsive components in $r(n)$ have been nearly eliminated in Fig. 3d. This is a result of quantizing these components with $R_O > 1$ bits/sample; $R_O = 4$ in this example. Since the impulsive components of the noise waveform $r(n)$ tend to occur predominantly during pitch-period onset, the system with non-instantaneous quantization can also be regarded as a form of "pitch-compensated" quantization.[8]

VI. SIGNAL-TO-NOISE RATIO RESULTS FOR R-BIT AND $(R + 1)$-BIT CODERS

Figures 7 and 8 compare the performance of the coders of Sections IV and V with that of conventional single-stage ADPCM.

The signal-to-noise ratios are averages over the entire length of a given utterance. The segmental s/n is obtained by obtaining the signal-

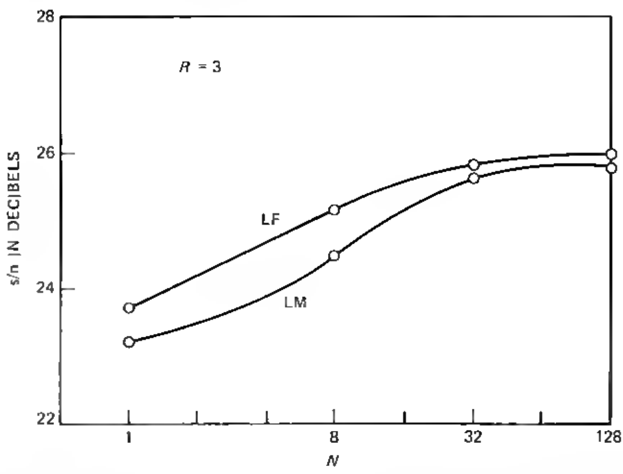


Fig. 6—Signal-to-noise ratio $(R + 1)$-bit ADPCM system with variable-rate quantization of reconstruction noise $r(n)$ from R-bit ADPCM. The signal-to-noise ratio reaches a value close to the maximum with a noise-buffer length of $N = 32$ (encoding delay of 4 ms). The gain over instantaneous noise quantization ($N = 1$) is in excess of 2 dB.

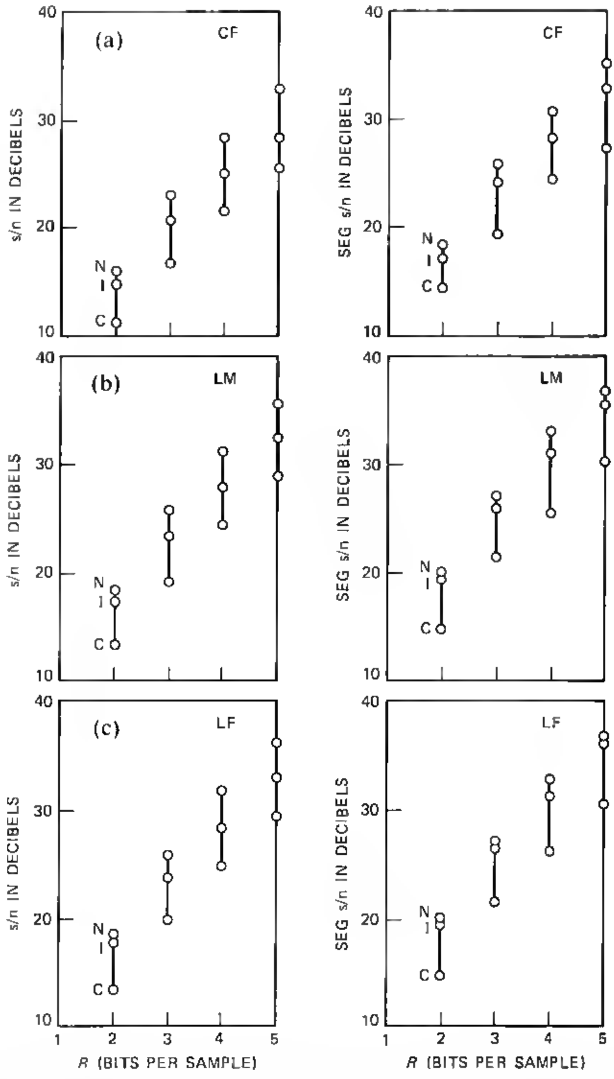


Fig. 7—Signal-to-noise ratio (s/n) and segmental signal-to-noise ratio (SEG s/n) in ADPCM systems, as a function of bit rate R of the basic coder in Fig. 1. For each value of R, there is an ordered set of three s/n or SEG s/n values. Plots in (a), (b), and (c) refer to speech inputs *CF*, *LM*, and *LF*.

to-noise ratio in dB for each 16-ms segment of an input, and by averaging such dB values over the entire length of a given utterance.

Figure 7 shows, for each bit rate R of the conventional ADPCM system (C), signal-to-noise ratio gains in $(R + 1)$-bit systems with *i*nstantaneous (I) and *n*on-instantaneous (N) quantization of $r(n)$, with a total of 32 bits of quantization in every 32-sample block of $r(n)$.

Note that except in the case of $R = 2$, the performance of the $(R + 1)$-bit system with instantaneous quantization (the middle point on each vertical bar) is very close to conventional $(R + 1)$-bit ADPCM (the lowest point on the next vertical bar to the right), with an s/n gap of no more than 1 dB. Note also that for $R > 2$, the $(R + 1)$-bit system with non-instantaneous quantization (the topmost point on each vertical bar) is always better than conventional $(R + 1)$-bit ADPCM, with an s/n gain of as much as 3 dB. The substantial gains at $R = 4$ and 5

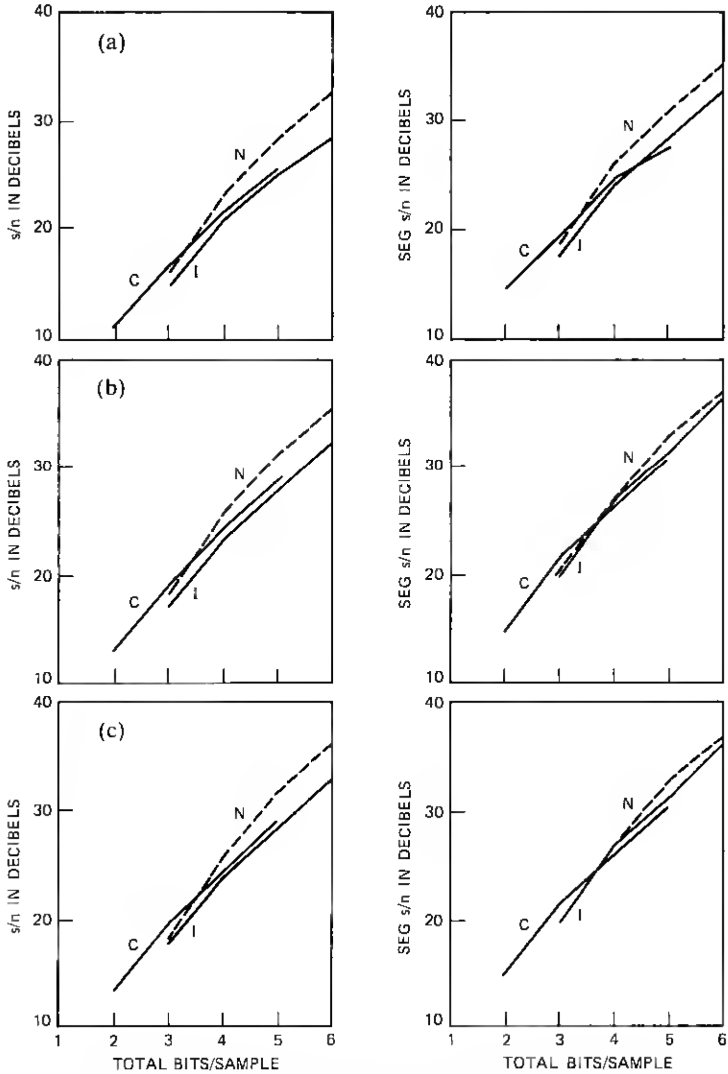


Fig. 8—Results of Fig. 7 replotted as a function of total bit rate.

may be due partly to the fact that the ADPCM quantizer in these cases is somewhat suboptimal; as R increases, optimal design of the 2^{R-1} step-size multipliers (Fig. 2) becomes increasingly difficult, and the s/n of the conventional ADPCM coder increases by less than the expected 6 dB per additional bit.

Figure 8 replots the results of Fig. 7, and compares the three coders discussed above, for given fixed values of total bit rate. Note once again that if the overall bit rate is at least 4 bits/sample, the $(R + 1)$-bit coder with instantaneous quantization is very close to conventional $(R + 1)$-bit ADPCM; while the $(R + 1)$-bit coder with non-instantaneous quantization is consistently better than $(R + 1)$-bit ADPCM.

VII. PERCEPTUAL EVALUATIONS OF THE CODERS OF SECTION IV AND V

Critical headphone listening reinforces the results suggested in Section VI. As expected, with $R = 2$, the outputs of 3-bit systems of Sections IV and V are both slightly worse than those of conventional 3-bit ADPCM. But with $R \geq 3$, even the output of the simpler $(R + 1)$-bit system (the system with instantaneous quantization) sounds extremely close to that of conventional $(R + 1)$-bit ADPCM. The very good perceptual performance of the instantaneous noise quantizer is very likely because much of its residual error (Fig. 3c) may be masked by the high-level speech activity in its temporal vicinity. In fact, the main motivation for the use of a non-instantaneous quantizer is not merely the increased performance with $(R + 1)$-bit coding, as demonstrated in Section VI, but also the fact that with more general $(R + R_n)$-bit coding $(R_n > 1)$, the performance of the instantaneous quantizer deteriorates rapidly, while that of the non-instantaneous quantizer maintains a 6-dB-per-bit behavior (Section VIII).

VIII. VARIABLE-RATE CODING WITH $R_n \geq 1$ BIT/SAMPLE

Sections IV through VII discussed the design and performance of a dual-rate system with $R_n = 0$ or 1, and a total bit rate of either R or $R + 1$ bits/sample. In this section, we consider a generalization to $R_n > 1$. Specifically, the average noise-coding bit rate R_n will range from 0 to 3, the ADPCM bit rate R will range from 2 to 5, and combinations of R and R_n will be such that the total bit rate $R + R_n$ will range from 2 to 6 bits/sample, the range used earlier in Fig. 8. We will note that the performance of an instantaneous noise-coding system deteriorates rapidly when $R_n > 1$, while that of a non-instantaneous noise-coding systems maintains an approximate 6-dB-per-additional-bit behavior.

8.1 Instantaneous noise coding

When $R_n = 1$, the recommended output levels for the APCM noise coder were $\pm 0.25\ \Delta(n)$. These levels are in fact centered in the ranges

0 to $\pm 0.5\ \Delta(n)$, the defining ranges for granular noise amplitude. As Fig. 4 shows, the design of the instantaneous quantizer is hardly affected by the occasional incidence of overload noise magnitudes much greater than $0.5\ \Delta(n)$. Generalizations to $R_n > 1$ therefore call for sets of 2^{R_n} APCM output levels that are uniformly spaced in the regions $-0.5\ \Delta(n)$ to $+0.5\ \Delta(n)$. For example, with $R_n = 2$ and 3, the output levels will be

$$R_n = 2\colon [\pm 0.125\Delta(n),\ \pm 0.375\Delta(n)]$$

and

$$R_n = 3\colon [\pm 0.0625\Delta(n),\ \pm 0.1875\Delta(n),\ \pm 0.3125\Delta(n),\ \pm 0.4375\Delta(n)]. \quad (9)$$

Experiments with $R_n = 2$ and 3 show that the above design is indeed nearly optimal for instantaneous coding. However, the performance of the instantaneous coder deteriorates badly as R_n increases, as we will see in Fig. 10. This is to be expected from the illustrative residual noise waveform of Fig. 3c, which shows that instantaneous coding is characterized by residual errors of very significant amplitude during periods of ADPCM overload. The situation does not improve with increasing R_n because the additional output levels that become available are simply used up for finer quantization in the granular noise region, shown in eq. (9).

8.2 Non-instantaneous noise coding

As we saw in the residual noise waveform of Fig. 3d for the example of average noise bit rate $R_n = 1$, non-instantaneous coding of the noise waveform can reduce the extent of granular noise as well as that of overload distortion in ADPCM coding. Slope-overload bursts are still visible in the residual noise waveform of Fig. 3d, but the waveform is much less impulsive than the original noise waveform of Fig. 3b. With $R_n > 1$, both the overload and granular components in Fig. 3d can be made increasingly smaller, provided that the bit allocation and quantizer design of Section V are properly generalized.

Recall that for an average noise bit rate of $R_n = 1$, the bit allocation (7) of Section V was as follows:

$$\begin{aligned} &R_O \text{ bits for } N_O \text{ overload noise samples} \\ &0 \text{ bits for } N_g = N_O(R_O - 1) \text{ low-amplitude noise samples} \qquad (10) \\ &1 \text{ bit for } N - N_O - N_g \text{ remaining noise samples.} \end{aligned}$$

The total number of bits is then N for every block of N samples, as in (7). As noted in (8), the above constraint also implies that $N_O \le N/R_O$. This condition has to be explicitly enforced even when the number of actual overload noise samples exceeds N/R_O for a chosen R_O. A simple

generalization of (10) that works very well with $R_n > 1$ is shown below:

$$\begin{aligned} &R_O + (R_n - 1) \text{ bits for } N_O \text{ overload noise samples} \\ &(R_n - 1) \text{ bits for } N_g = N_O(R_O - 1) \text{ low-amplitude noise samples} \\ &R_n \text{ bits for } N - N_O - N_g \text{ remaining noise samples.} \end{aligned} \tag{11}$$

The total number of bits is now NR_n for every block of N samples. Furthermore, (11) is a straightforward generalization of (10); and as in the case of (10), it requires no transmission of any side information for bit-allocation purposes, but only encoding and decoding delays in the order of $N = 32$ (4 ms, assuming 8-kHz samples).

Critical to the success of the bit-allocation algorithm (11) is a proper design of individual quantization characteristics. Unlike the instantaneous design of (9), the variable-bit allocations in (11) permit finer quantizer resolutions both in the overload range, $|r(n)| > 0.5\Delta(n)$, and in the granular noise range, $|r(n)| < 0.5\Delta(n)$. A systematic way to design these quantizers is to start with the designs in Section V (for a given R, and for an average-noise-bit rate of $R_n = 1$). Recall that each such design involves a set of three characteristics, for 0-bit, 1-bit, and $R_O = R$-bit quantization, as in (10). As the value of R_n increases, each of these sets evolves into corresponding sets of three characteristics, for $(R_n - 1)$-bit, R_n-bit, and $(R_O + R_n - 1)$-bit quantization, as in (11). Resolutions improve by a factor of two for each stage of increase of R_n, and this improvement benefits the overload as well as granular regions of coding noise. Figure 9 illustrates the quantizer evolution for the example of $R = 3$ and $R_n = 1$ and 2. The illustration includes only one of the set of three quantizers involved in the coding process. This is the R_O-bit characteristic (Fig. 9a, which is the same as Fig. 5b) used for quantizing the N_O overload noise samples in the $R_n = 1$ system. When $R_n = 2$, the above R_O-bit (in this case, 3-bit) characteristic evolves into a $R_O + R_n - 1 = 4$-bit characteristic (Fig. 9b).

Figure 10 shows the benefits of increasing R_n in a non-instantaneous noise-coding system, for the example of $R = 4$ and for average-noise-coding rates of $R_n = 1$, 2, and 3 bits/sample. All error waveforms are magnified by a factor of 50. The waveform in (b) is the same as that in Fig. 3d, but is magnified by a factor of 5. In Fig. 10 we see a significant reduction in residual noise level for each stage of increasing R_n. We will note presently that the improvement is very close to 6 dB per additional bit in R_n.

8.3 Signal-to-noise ratios

Figures 11a and b show s/n and segmental s/n results for explicit noise-coding systems with $R_n \geq 1$. The range of total bit rate $R + R_n$ is 2 to 6, the same as that in Fig. 8. The solid curves show the

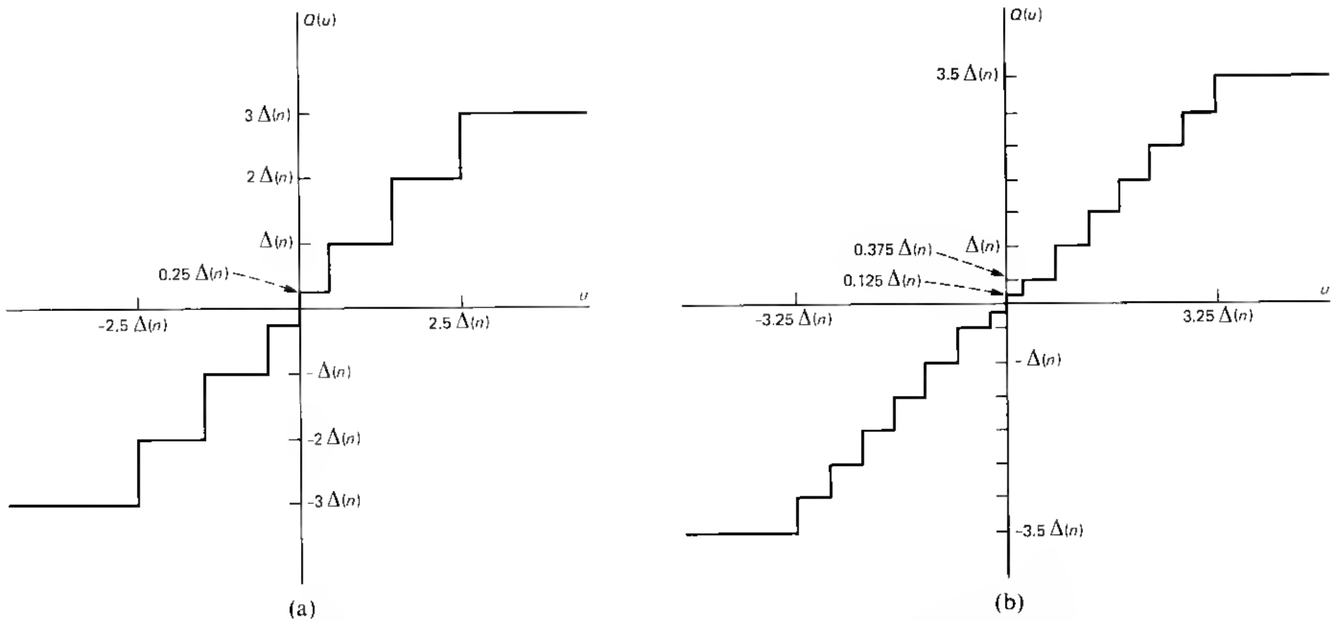


Fig. 9—Examples of quantization characteristics used for overload noise samples $r_O(n)$ in the non-instantaneous coding of 3-bit ADPCM noise with an average rate of (a) $R_n = 1$ bit/sample and (b) $R_n = 2$ bits/sample.

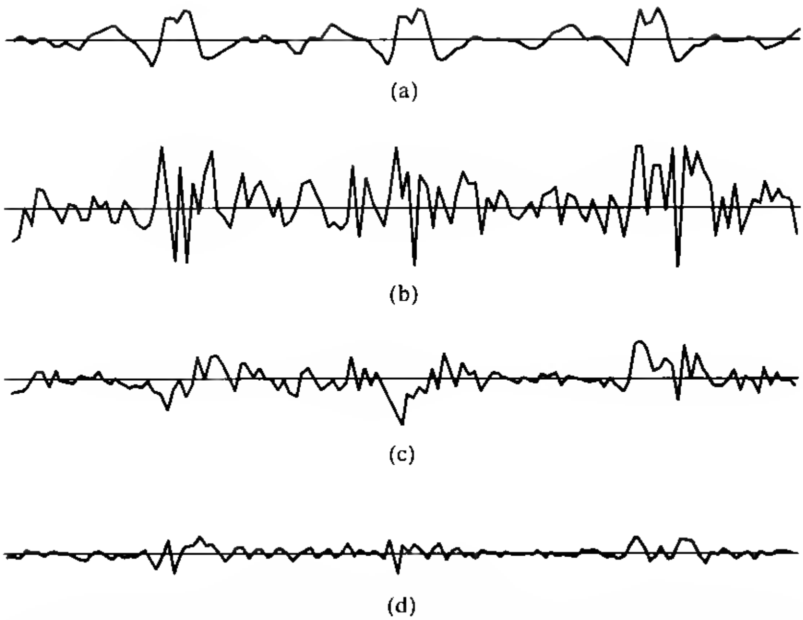


Fig. 10—(a) Input speech $x(n)$ and reconstruction error waveforms in variable-rate coding with 4-bit ADPCM, non-instantaneous noise coding and average-noise-coding rates of (b) $R_n = 1$, (c) $R_n = 2$, and (d) $R_n = 3$ bits/sample. All error waveforms are magnified by a factor of 50. The waveform in (b) is the same as that in Figure 3(d), but magnified by a factor of 5.

performance of conventional ADPCM. The circles labelled 3 show the performance of a variable-rate system based on instantaneous coding of ADPCM noise, for the example of $R = 3$. We can see that with $R_n > 1$, the s/n performance of the instantaneous coding system deteriorates fairly rapidly, compared with that of $(R + R_n)$-bit ADPCM, with increasing bit rate, but its segmental s/n performance is competitive with that of conventional ADPCM at all bit rates. Non-instantaneous coding systems, on the other hand, maintain a 6-dB per additional bit behavior, provided only that $R > 2$. This is shown by the sets of solid black dots labelled 3, 4, and 5. The performances of these systems also exceed that of conventional ADPCM at any given total bit rate, a result already noted in Section VI for the special case of $R_n = 1$.

IX. EFFECTS OF TRANSMISSION ERRORS

Bit errors in transmission can affect the noise-coding system in two ways: they may produce effects attributable to errors in the transmission of the bits from the basic DPCM coder, and effects attributable to errors in the bits from the noise coder. Effects of both types are expected to be more severe in the case of the non-instantaneous coder.

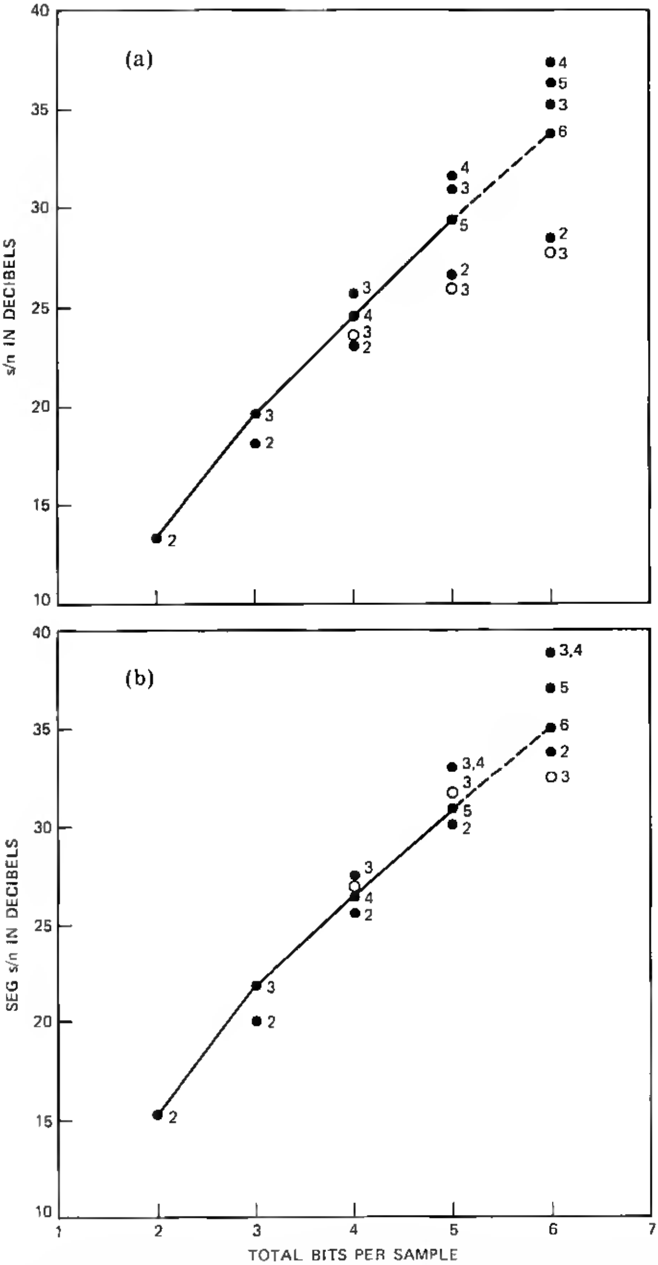


Fig. 11—(a) Signal-to-noise ratio (s/n) and (b) segmental signal-to-noise ratio (SEG s/n) in ADPCM systems, as a function of total bit rate. The solid curves refer to conventional ADPCM ($R_n = 0$) and the black dots refer to non-instantaneously quantized noise-coding systems with $R = 2$, 3, 4, and 5 and $R_n \geq 1$ bits/sample. The circles refer to an instantaneously quantized noise-coding system with $R = 3$, and $R_n = 1$, 2, and 3 bits/sample.

The greater error sensitivity of this system is due first to the presence of quantizers with larger step sizes, and hence, proportionally larger channel-error effects. More important, the increased error sensitivity of the non-instantaneous system is a result of the variable-bit-allocation algorithm, which will be miscalculated at the receiver if one or more bits from the basic ADPCM coders are received in error. Errors of this type do not propagate beyond a given N-sample block, but their effects can be severe enough to warrant the complete disabling of the noise-coding part of the system when errors are detected. A simple example of an error-detection system is one where the odd-even parity of the number N_O of overload samples is explicitly transmitted to the receiver. A change in the parity of N_O, as computed at the receiver, is a good detector of perceptually significant single-bit errors in the given block. The single bit needed to transmit the parity information, or the multiplicity of bits needed to transmit the information in an error-protected format, can be incorporated in the coder output by a bit-stealing procedure based on increasing the number of zero-bit noise samples from N_g to an appropriately greater number.

Irrespective of the noise-coding method and the procedures that may be used to protect the noise-coding system from errors, the basic ADPCM coder can be made error-robust, at least for independent error rates of up to 10^{-3}, by using robust adaptive-quantizer algorithms such as the leaky-adaptation algorithm in Ref. 9.

X. CONCLUSIONS

We have demonstrated simple systems for variable-rate, embedded ADPCM coding of speech based on explicit coding of reconstruction noise. These systems do not require the transmission of any side information other than what is already available in a conventional ADPCM decoder. The simpler of two systems proposed in this paper uses instantaneous coding of the noise, and provides a performance very close to that of the conventional ADPCM at any given value of total bit rate $R_T = R + R_n$, for the simple but non-trivial case of dual-rate operation ($R_n = 0$ or 1 bit/sample). But its s/n performance deteriorates significantly with more widely variable operation ($R_n > 1$ bits/sample). The more complex system uses non-instantaneous noise-coding, with coding and decoding delays in the order of 4 ms to realize positive gains over conventional ADPCM at any given total bit rate $R + R_n$ bits/sample. The performance of this system has been demonstrated for $R_n = 0$, 1, 2, and 3 bits/sample, and for $R = 2$, 3, 4, and 5 bits/sample. The system with non-instantaneous noise coding can also be regarded as an $(R + R_n)$-bit ADPCM coder with a quantizing system that is better than conventional adaptive quantization with a one-word memory.

REFERENCES

1. J. Das and P. K. Chatterjee, "Optimized Δ − Δ Modulation System," Elect. Lett., *3*, No. 6 (June 1967), pp. 286–7.
2. R. Steele, "Peak Signal-to-Noise Formulas for Multistage Delta Modulation with *RC*-Shaped Caussian Input Signals," B.S.T.J., *61*, No. 3 (March 1982), pp. 347–62.
3. T. Bially, B. Cold, and S. Seneff, "A technique for adaptive voice flow control in integrated packet networks," IEEE Trans. Commun. (March 1980), pp. 325–33.
4. Y.-C. Ching, "Differential Pulse Code Communications System Having Dual Quantization Schemes," U.S. Patent 3,781,685, December 25, 1973.
5. D. J. Goodman, "Embedded DPCM for Variable Bit Rate Transmission," IEEE Trans. Commun. (July 1980), pp. 1040–46.
6. P. Cummiskey, N. S. Jayant, and J. L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech," B.S.T.J., *52*, No. 7 (September 1973), pp. 1105–18.
7. N. S. Jayant, "Adaptive Quantization With a One-Word Memory," B.S.T.J., *52*, No. 7 (September 1973), pp. 1119–44.
8. D. L. Cohn and J. L. Melsa, "The Residual Encoder—An Improved ADPCM System for Speech Digitization," IEEE Trans. Commun. (September 1975), pp. 935–41.
9. D. J. Goodman and R. M. Wilkinson, "A Robust Adaptive Quantizer," IEEE Trans. Commun. (November 1975), pp. 1362–5.